Central limit theorem for fluctuations in the high temperature region of the Sherrington-Kirkpatrick spin glass model

Francesco Guerra[*]

Dipartimento di Fisica, Università di Roma 'La Sapienza'
and INFN, Sezione di Roma, Piazzale A. Moro 2, 00185 Roma, Italy

Fabio Lucio Toninelli[†]

Scuola Normale Superiore, Piazza dei Cavalieri 7, 56126 Pisa, Italy
and Istituto Nazionale di Fisica Nucleare, Sezione di Pisa

February 1, 2008

Abstract

In a region above the Almeida-Thouless line, where we are able to control the thermodynamic limit of the Sherrington-Kirkpatrick model and to prove replica symmetry, we show that the fluctuations of the overlaps and of the free energy are Gaussian, on the scale $1/\sqrt{N}$, for N large. The method we employ is based on the idea, we recently developed, of introducing quadratic coupling between two replicas. The proof makes use of the cavity equations and of concentration of measure inequalities for the free energy.

[*] e-mail: `francesco.guerra@roma1.infn.it`
[†] e-mail: `f.toninelli@sns.it`

1 Introduction

We consider the mean field spin glass model introduced by Sherrington and Kirkpatrick in [1], [2], in the regime of high temperature or, equivalently, of large magnetic field. On physical grounds, it is known that in this region the replica symmetric solution holds, as shown for example in [3], and references quoted there. However, due to the very large fluctuations present in the model, it is difficult to give a mathematically rigorous description of this region. Rigorous works on this subject include [4], [5], [6], [7], [8]. For other rigorous results concerning the model, we refer to [9], [10], [11], [12], [13].

The method developed in [8] by Michel Talagrand is particularly interesting. The starting point is the very deep physical idea that the phenomenon of replica symmetry breaking can be understood by studying the properties of the model under the application of auxiliary interactions, which explicitly break replica symmetry. In [8], this idea is employed to prove that the replica symmetric solution holds in a region, which (*probably*) coincides with that found in the theoretical physics literature [3], i.e., up to the Almeida-Thouless critical line.

Recently [14] we proposed a different strategy, which consists in coupling two replicas of the system by means of a term proportional to the square of the deviation of the overlap from its replica symmetric value. In this way, we proved that replica symmetry holds in a region above the Almeida-Thouless line. In the same region, we obtained a control of the two-replica system, provided that the coupling parameter is small enough, and we showed that the fluctuations of the overlap are at most of order $1/\sqrt{N}$. In the present paper we prove that, in the same region of parameters, the fluctuations of overlaps and free energy, when suitably rescaled, have a Gaussian distribution when $N \to \infty$. The main ingredients of the proof are the control of the thermodynamic limit obtained in [14] and concentration of measure techniques inspired by Talagrand's works. Then, by means of the cavity method, one can write self-consistent linear equations for the characteristic functions of the fluctuation variables, which can be easily solved.

Previous results concerning limit theorems for fluctuations in the high temperature region of mean field spin glass models include [4], [6], [15], [9], [16].

This work is organized as follows: In Section 2, we recall the main definitions of the model and introduce the overlap distribution structure. In Section 3, we state the main results. Two useful tools, i.e., exponential inequalities and the cavity method, are briefly outlined in Sections 4 and 5. In Sections 6 and 7, we prove the central limit theorem for overlap and free energy fluctuations, respectively. Finally, Section 8 is dedicated to a short outlook about open problems and further developments.

2 The model

The generic configuration of the Sherrington-Kirkpatrick (SK) model is determined by the N Ising variables $\sigma_i = \pm 1$, $i = 1, 2, \ldots, N$, and the Hamiltonian is

$$H_N(\sigma, h; J) = -\frac{1}{\sqrt{N}} \sum_{(i,j)} J_{ij}\sigma_i\sigma_j - h \sum_i \sigma_i, \tag{1}$$

where the sum $\sum_{(i,j)}$ runs over all the $N(N-1)/2$ distinct couples of sites. The J_{ij}'s (quenched noise) are independent centered unit Gaussian variables $\mathcal{N}(0,1)$. The first term in (1) is a long range random two body interaction, while the second represents the interaction with a fixed external magnetic field h. For a given temperature $1/\beta$ we can introduce the disorder dependent partition function

$$Z_N(\beta, h; J) = \sum_{\{\sigma\}} \exp(-\beta H_N(\sigma, h; J)) \tag{2}$$

and the auxiliary function

$$\alpha_N(\beta, h) = \frac{1}{N} E \ln Z_N(\beta, h; J), \tag{3}$$

where E denotes the average with respect to the external noise J. Note that $\alpha_N(\beta, h)$ is the quenched average of the free energy per spin, apart from the multiplicative factor $-1/\beta$.

For later convenience, it is useful to generalize the model (1) by introducing a quenched random external magnetic field, which at every site is an independent Gaussian variable of strength $x > 0$. In other words, the Boltzmann factor of the system becomes

$$\exp\left(\sqrt{\frac{t}{N}} \sum_{(i,j)} J_{ij}\sigma_i\sigma_j + \sum_i (\beta h + \sqrt{x}J_i)\sigma_i\right), \tag{4}$$

where J_i are i.i.d. $\mathcal{N}(0,1)$ random variables, independent of the J_{ij}'s. We let $t = \beta^2$ in the two-body term. In the following, we always regard the system as depending on the parameters $t, x, \beta h$. In analogy with Eqs. (2), (3), we define the disorder dependent partition function $Z_N(t, x, h; J)$ and the auxiliary function

$$\alpha_N(t, x, h) = \frac{1}{N} E \ln Z_N(t, x, h; J).$$

Here, E denotes averaging with respect to J_{ij} and J_i. For simplicity of notations, here and in the following we write the argument h instead of βh.

Let us consider a countably infinite number of independent copies (replicas) of the system, whose spin variables σ_i^a are distributed, for fixed J, according to the product state

$$\Omega_J = \Omega_J^1 \Omega_J^2 \ldots,$$

where $\Omega_J \equiv \Omega_J^{N,t,x,h}$ denotes the Gibbs state associated to the Boltzmann factor (4). Each replica is subject to the same quenched noise. The "real replica" approach has already been exploited in a number of papers [17], [18], [19], [20].

The overlap between two replicas a, b is defined as

$$q_{ab}(\sigma^a, \sigma^b) = \frac{1}{N} \sum_i \sigma_i^a \sigma_i^b,$$

with the obvious bounds

$$-1 \leq q_{ab} \leq 1.$$

For a generic smooth function F of the overlaps, we define the $\langle . \rangle$ average as

$$\langle F(q_{12}, q_{13}, \ldots) \rangle = E\Omega_J(F(q_{12}, q_{13}, \ldots)).$$

Note that the average over disorder introduces correlations between different groups of replicas, which would be independent under the Boltzmann average Ω_J. For example,

$$\Omega_J(q_{12}q_{34}) = \Omega_J(q_{12})\Omega_J(q_{34})$$

but

$$\langle q_{12}q_{34} \rangle \neq \langle q_{12} \rangle \langle q_{34} \rangle.$$

3   The high temperature region and the main results

In this Section, we recall the results of [14] and state limit theorems for fluctuations, in the region where we prove that replica symmetry holds, i.e.,

$$\lim_{N \to \infty} \alpha_N(t, x, h) = \bar{\alpha}(t, x, h).$$

$\bar{\alpha}(t, x, h)$ is the replica-symmetric free energy [1], [2]

$$\bar{\alpha}(t, x, h) = \ln 2 + \int \ln \cosh(\beta h + z\sqrt{t\,\bar{q} + x})\, d\mu(z) + \frac{t}{4}(1 - \bar{q})^2,$$

$\bar{q}$ is the Sherrington-Kirkpatrick order parameter, defined as the unique [17] solution of

$$\bar{q} = \bar{q}(t, x, h) = \int \tanh^2(\beta h + z\sqrt{t\,\bar{q} + x})\, d\mu(z)$$

and $d\mu(z)$ is the centered unit Gaussian measure.

In [14] we proved the following: Consider the auxiliary function $\tilde{\alpha}_N$, dependent on the parameter $\lambda \geq 0$

$$\tilde{\alpha}_N(t, x, h; \lambda) = \alpha_N(t, x, h) + \frac{1}{2N} E \ln \Omega_{t,x,h}\left(e^{N\frac{\lambda}{2}(q_{12} - \bar{q})^2}\right)$$

4

and the trajectory in the (t, x) plane

$$\Gamma = (t', x_{t'}) \equiv (t', x + \bar{q}(t - t')) \equiv (t', x_0 - \bar{q}\, t'), \qquad 0 \leq t' \leq t \qquad (5)$$

where $x_0 = x + \bar{q}\, t$ and $\bar{q} = \bar{q}(t, x, h) = \bar{q}(t', x_{t'}, h)$. Notice that $\tilde{\alpha}_N$ equals α_N for $\lambda = 0$. Given x_0, h there exists a value $t_c(x_0, h)$, such that

$$|\bar{\alpha}(t', x_{t'}, h) - \tilde{\alpha}_N(t', x_{t'}, h; \lambda)| \leq \frac{k}{N} \qquad (6)$$

for some constant k, uniformly in the triangular region

$$0 \leq t' + \lambda \leq \bar{t} < t_c(x_0, h). \qquad (7)$$

In the same region, the overlaps self-average around the value $\bar{q}$:

$$\langle (q_{ab} - \bar{q})^2 \rangle \leq \frac{k}{N}. \qquad (8)$$

The critical value $t_c(x_0, h)$ is determined in the following way [14]: Let

$$\Delta(x_0, h, \lambda_0) \equiv \frac{1}{2} \max_{\rho \in \mathbb{R}} \left(\int \ln(\cosh \rho + \tanh^2(\beta h + z\sqrt{x_0})\sinh \rho)d\mu(z) - \rho\,\bar{q} - \frac{\rho^2}{2\lambda_0} \right),$$

where $\lambda_0 \geq 0$. Then, we define $t_c(x_0, h)$ such that, for any $\lambda_0 \leq t_c(x_0, h)$, one has

$$\Delta(x_0, h, \lambda_0) = 0.$$

In the case of vanishing external field $x=h=0$, then also $x_0=\bar{q}=0$ and $t_c = 1$, the correct critical value. As discussed in [14], the region defined by (7) falls short of the Almeida-Thouless line, which is the expected critical line.

In this paper, we investigate more precisely the behavior of fluctuations of physical quantities around the replica symmetric value. First of all, we give a central limit-type theorem for the rescaled overlaps

$$\xi_{ab}^N = \sqrt{N}(q_{ab} - \bar{q}),$$

showing that they behave as centered Gaussian variables characterized by a non-diagonal correlation matrix. Notice that, thanks to (8), one has the following bound for the second moment of the rescaled overlap fluctuations:

$$\langle (\xi_{ab}^N)^2 \rangle \leq k. \qquad (9)$$

Theorem 1. *If $t < t_c(x_0, h)$, the rescaled overlaps ξ_{ab}^N tend in distribution, for $N \to \infty$, to jointly Gaussian variables ξ_{ab}, with covariances*

$$\langle \xi_{ab}^2 \rangle = A(t, x, h)$$
$$\langle \xi_{ab}\xi_{ac} \rangle = B(t, x, h)$$
$$\langle \xi_{ab}\xi_{cd} \rangle = C(t, x, h),$$

5

where $b \neq c$, $c \neq a, b$ and $d \neq a, b$. A, B and C are explicitly given by

$$A(t, x, h) = (1 + 2R + 4R^2)Y + c_0 R^2 \qquad (10)$$
$$B(t, x, h) = (1 + 4R)RY + c_0 R^2 \qquad (11)$$
$$C(t, x, h) = 4R^2 Y + c_0 R^2, \qquad (12)$$

where

$$Y(t, x, h) = \frac{1}{Y_0^{-1} - t}$$
$$R(t, x, h) = \frac{d_0}{Y_0^{-1} + 2d_0 - t}$$

and $Y_0(x_0, h)$, $c_0(x_0, h)$ and $d_0(x_0, h)$ are chosen in such a way that A, B, C satisfy the initial conditions

$$A(0, x_0, h) = 1 - \bar{q}^2$$
$$B(0, x_0, h) = \bar{q} - \bar{q}^2$$
$$C(0, x_0, h) = \int \tanh^4(z\sqrt{x_0} + \beta h) \, d\mu(z) - \bar{q}^2.$$

In particular, one has

$$Y_0 = \int \cosh^{-4}(z\sqrt{t\bar{q} + x} + \beta h) \, d\mu(z) \qquad (13)$$

Recently, an analogous result was proved independently by Talagrand [16], who computed the $N \to \infty$ limit for all moments of the ξ variables.

The expressions for A, B and C were first given by Guerra in [17]. For $h = x = 0$, the limit Gaussian variables are not correlated and have variance $1/(1-t)$, which is a well known result [4], [6].

Let us consider now free energy fluctuations. Aizenman, Lebowitz and Ruelle [4] proved that in the case of zero external field and $t < 1$, the variable

$$\ln Z_N - \ln E Z_N$$

tends to a shifted Gaussian random variable whose variance diverges at $t = 1$. In the general case the situation is quite different and the following theorem holds:

Theorem 2. *Let*

$$\hat{f}_N(t, x, h; J) \equiv \sqrt{N} \left(\frac{\ln Z_N(t, x, h; J)}{N} - \bar{\alpha}(t, x, h) \right).$$

If $t < t_c(x_0, h)$ then

$$\hat{f}_N(t, x, h; J) \xrightarrow{d} \mathcal{N}(0, \sigma^2(t, x, h)),$$

where

$$\sigma^2(t, x, h) = Var\left(\ln\cosh(z\sqrt{t\,\bar{q} + x} + \beta h)\right) - \frac{\bar{q}^2 t}{2}$$

Here, $Var(.)$ denotes the variance of a random variable and $z = \mathcal{N}(0,1)$.

Notice that fluctuations of the extensive free energy $\ln Z_N$ are of order 1 at zero external field and of order $\sqrt{N}$ otherwise.

4 Exponential suppression of overlap fluctuations

General arguments based on concentration of measure [9], [21], [22] show that the fluctuations of the free energy $1/N \ln Z_N$ around its mean value α_N are exponentially suppressed as N grows. Indeed, one has the following [9]

Theorem 3. *For any $u > 0$,*

$$P\left(\left|\frac{1}{N}\log Z_N(t, x, h; J) - \alpha_N(t, x, h)\right| \geq u\right) \leq \exp(-NKu^2), \qquad (14)$$

where

$$K = \frac{1}{t + 2x}.$$

This, in connection with the results of [14], allows to obtain a strong control on the fluctuations of the overlaps (we learned this nice argument in [8]): First of all, the same argument leading to Theorem 3 shows that

$$P\left(\left|\frac{1}{2N}\ln\Omega_{t,x,h}\left(e^{N\frac{\lambda}{2}(q_{12}-\bar{q})^2}\right) - \tilde{\alpha}_N(t, x, h, \lambda) + \alpha_N(t, x, h)\right| \geq 2u\right) \leq \exp(-NKu^2).$$

Therefore, thanks to Eq. (6), with probability at least $1 - \exp(-NKu^2)$ one has

$$\Omega_{t,x,h}\left(e^{\frac{\lambda}{2}N(q_{12}-\bar{q})^2}\right) \leq e^{4Nu+2C}$$

for $\lambda \leq \bar{\lambda} < t_c(x_0, h) - t$. Then, by Tchebyshev's inequality

$$\Omega_{t,x,h}\left(\chi_{\{|q_{12}-\bar{q}|\geq v\}}\right) \leq e^{-\frac{\bar{\lambda}}{2}Nv^2}\Omega_{t,x,h}\left(e^{\frac{\bar{\lambda}}{2}N(q_{12}-\bar{q})^2}\right) \leq e^{N(4u-\frac{\bar{\lambda}}{2}v^2)+2C}$$

and, choosing $u = \bar{\lambda}v^2/16$, one has

$$\Omega_{t,x,h}\left(\chi_{\{|q_{12}-\bar{q}|\geq v\}}\right) \leq e^{-N\frac{v^2\bar{\lambda}}{4}+2C},$$

The estimate we are looking for easily follows:

$$E\,\Omega_{t,x,h}\left(\chi_{\{|q_{12}-\bar{q}|\geq v\}}\right) \leq e^{-N\frac{v^2\bar{\lambda}}{4}+2C} + e^{-NK\frac{\bar{\lambda}^2v^4}{256}}. \qquad (15)$$

Of course, this is much more than just self-averaging of the overlaps.

5 The cavity method

The cavity method allows to express thermal averages of quantities defined on the N-spin system as functions of averages on the system with $N - 1$ spins, at a slightly different temperature. This method has been widely applied both in the theoretical physics literature [3] and in the mathematical physics one (see, for instance, [5], [16], [10], [18]).

Introduce the following definitions:

$$t' = t\left(1 - N^{-1}\right)$$
$$\sigma^a = (\eta^a, \epsilon^a), \qquad \eta^a \in \{-1, 1\}^{N-1}, \qquad \epsilon^a = \sigma_N^a = \pm 1$$
$$J = J_N$$
$$g_i = J_{N\,i}, \qquad i = 1, \ldots, N - 1$$
$$\Omega'(.) = \Omega_{N-1}^{t'}(.)$$

The cavity equations consist in the identity

$$\Omega_N^{t,x,h}\left(f(\sigma^1, \ldots, \sigma^k)\right) = \frac{\Omega'\left(Av\, f(\eta^1, \epsilon^1, \ldots, \eta^k, \epsilon^k)\Psi^{(k)}\right)}{\Omega'\left(Av\,\Psi^{(k)}\right)}, \qquad (16)$$

where Av denotes the average over the spin variables ϵ^a and

$$\Psi^{(k)} \equiv \exp \sum_{a=1}^{k} \epsilon^a \left(\sqrt{t/N}\,g\,\eta^a + \sqrt{x}J + \beta h\right). \qquad (17)$$

$g\,\eta^a$ denotes the scalar product $\sum_{i=1}^{N-1} g_i \eta_i^a$.

6 Limit theorem for overlap fluctuations

To prove Theorem 1, it suffices [23] to show that for any integer s, the characteristic function

$$\phi_N^t(u) = \left\langle \exp i\,u\,\xi^N \right\rangle \equiv \left\langle \exp i \sum_{(a,b)} u_{ab}\,\xi_{ab}^N \right\rangle, \qquad 1 \le a < b \le s$$

converges for $N \to \infty$ to

$$\phi^t(u) = \exp\left\{-\frac{1}{2}\left(\hat{L}u, u\right)\right\}, \qquad (18)$$

where $(.\,,.)$ denotes scalar product and $\hat{L}$ is the $s(s-1)/2 \times s(s-1)/2$ dimensional matrix of elements

$$L_{(ab),(ab)} = A(t, x, h)$$
$$L_{(ab),(ac)} = B(t, x, h)$$
$$L_{(ab),(cd)} = C(t, x, h).$$

The idea of the proof is to obtain a set of closed linear differential equations for $\phi_N^t(u)$, which determine uniquely the solution as (18), for $N \to \infty$. Some of the calculations involved in the proof are quite long, although straightforward, and are therefore just sketched.

First of all, we explain how the cavity equations (16), (17) can be simplified in the region where (15) holds. Following [10], we introduce some notations, letting $\Omega(.) \equiv \Omega_N^{t,x,h}(.)$ and $\Omega'(.) \equiv \Omega_{N-1}^{t',x,h}(.)$. Moreover, we define

$$
\begin{aligned}
b &= \Omega'(\eta) \in \mathbb{R}^{N-1} \\
\dot{\eta}^a &= \eta^a - b \\
X &= \sqrt{t/N}\, g\, b + \sqrt{x} J + \beta h \\
\Psi_0^{(k)} &= \exp(X \sum_{a=1}^{k} \epsilon^a) \\
f(\sigma^1, \ldots, \sigma^k) &= f(\eta^1, \epsilon^1, \ldots, \eta^k, \epsilon^k).
\end{aligned}
$$

Then, the following holds [10]:

Theorem 4.

$$
\begin{aligned}
E\Omega\left(f(\sigma^1, \ldots, \sigma^k)\right) &= E\frac{1}{\cosh^k X}\Omega'\left(Av\, f\, \Psi_0^{(k)}\right) &\quad (19) \\
&\quad + t\, E\frac{1}{\cosh^k X}\Omega'\left(Avf\, \Psi_0^{(k)} \sum_{1 \le a < c \le k} \epsilon^a \epsilon^c \frac{\dot{\eta}^a \dot{\eta}^c}{N}\right) &\quad (20) \\
&\quad + t\, E\frac{1}{\cosh^k X}\Omega'\left(Avf\, \Psi_0^{(k)} \sum_{1 \le a \ne c \le k} \epsilon^a \epsilon^c \frac{\dot{\eta}^c b}{N}\right) &\quad (21) \\
&\quad - k\, t\, E\frac{\tanh X}{\cosh^k X}\Omega'\left(Avf\, \Psi_0^{(k)} \sum_{a=1}^{k} \epsilon^a \frac{\dot{\eta}^a b}{N}\right) + S &\quad (22)
\end{aligned}
$$

and the "error term" S can be estimated as

$$
|S| \le w_k(t, x, h) E\Omega'\left(Av|f|\left(\sum_{a=1}^{k+1}\left(\frac{\dot{\eta}^a b}{N}\right)^2 + \sum_{1 \le a < c \le k+2}\left(\frac{\dot{\eta}^a \dot{\eta}^c}{N}\right)^2\right)\right),
$$

where w is a smooth function of its arguments, independent of N.

Note that, with respect to Theorem 3.2 in [10], the last sum in the r.h.s. is performed on $a < c$ instead of $a \le c$. However, the proof of Theorem 4 proceeds exactly as in [10].

Theorem 4 is a sort of Taylor expansion of the cavity equations around $\eta^a = b$. This turns out to be particularly useful in the region where Eq. (15) holds, since

9

in this case $\dot{\eta}_a$ is small with large probability, and S vanishes for $N \to \infty$, as we explain in the following.

In order to prove Theorem 1, we first exploit symmetry between sites to write

$$\partial_{u_{rr'}} \phi_N^t(u) = i \left\langle \xi_{rr'}^N e^{i u \xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^r \sigma_N^{r'} - \bar{q}) e^{i u \xi^N} \right\rangle \tag{23}$$

$$\varphi_N^{(a)\,t}(u) \equiv i \left\langle \xi_{a,s+1}^N e^{i u \xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^a \sigma_N^{s+1} - \bar{q}) e^{i u \xi^N} \right\rangle \tag{24}$$

$$\psi_N^t(u) \equiv i \left\langle \xi_{s+1,s+2}^N e^{i u \xi^N} \right\rangle = i\sqrt{N} \left\langle (\sigma_N^{s+1} \sigma_N^{s+2} - \bar{q}) e^{i u \xi^N} \right\rangle, \tag{25}$$

and then employ the cavity equations to express these quantities as functions of ϕ, φ, ψ themselves. For instance, apply Theorem 4 to the r.h.s. of Eq. (23) and consider the term arising from (19). After averaging on the dichotomic variables ϵ, one is left with

$$i\left(\sqrt{N} - i \sum_{(a,b)} u_{ab}\,\bar{q}\right) E\left\{(\tanh^2 X - \bar{q})\Omega' \exp\left(i\,u'\,\xi^{N-1}\right)\right\} + \tag{26}$$

$$- u_{rr'} E\left\{(1 - \bar{q}\tanh^2 X)\Omega' \exp\left(i\,u'\,\xi^{N-1}\right)\right\}$$

$$-(1 - \bar{q}) \sum_{a \neq r,r'} (u_{ar} + u_{ar'}) E\left\{\tanh^2 X\,\Omega' \exp\left(i\,u'\,\xi^{N-1}\right)\right\}$$

$$- \sum_{(c,d)\,c,d \neq r,r'} u_{cd}\, E\left\{\tanh^2 X \left(\tanh^2 X - \bar{q}\right)\Omega' \exp\left(i\,u'\,\xi^{N-1}\right)\right\} + o(1),$$

where $u' = u\sqrt{1 - 1/N}$. The term $o(1)$ arises when $\exp\left(iu/\sqrt{N}\right)$ is expanded around $u=0$ and the terms of order u^2 or higher are neglected. Indeed, one has

$$\left|\sqrt{N} \left\langle (\sigma_N^r \sigma_N^{r'} - \bar{q}) e^{iu'\xi^{N-1}} \left(e^{i\sum_{(a,b)} \frac{u_{ab}}{\sqrt{N}}(\sigma_N^a \sigma_N^b - \bar{q})} - 1 - i\sum_{(a,b)} \frac{u_{ab}}{\sqrt{N}}(\sigma_N^a \sigma_N^b - \bar{q})\right) \right\rangle\right|$$

$$\leq 2\sqrt{N} \left|e^{i\sum_{(a,b)} \frac{u_{ab}}{\sqrt{N}}(\sigma_N^a \sigma_N^b - \bar{q})} - 1 - i\sum_{(a,b)} \frac{u_{ab}}{\sqrt{N}}(\sigma_N^a \sigma_N^b - \bar{q})\right| = O(N^{-1/2}).$$

Now, rewrite E as $E\,E_g$, where E_g denotes the average only with respect to the random variables J and $g_i, i = 1, \ldots, N-1$, and notice that J, g_i do not appear in the thermal average Ω'. Computation of $E_g(\ldots)$ would be simpler if, instead of X, there were

$$\bar{X} \equiv \sqrt{t/N}\,g\,\bar{b} + \sqrt{x}J + \beta h,$$

where

$$\bar{b} \equiv \frac{b}{||b||}\sqrt{N\bar{q}}.$$

Of course, one has

$$\bar{X} \overset{d}{=} z\sqrt{t\,\bar{q} + x} + \beta h,$$

10

where z is a standard unit Gaussian variable and equality holds in distribution so that, for instance,

$$E_g \tanh^2 \bar{X} = \bar{q}.$$

The idea is, therefore, to expand around $X = \bar{X}$. As a preliminary fact, notice that the second moment of the random variable $(b - \bar{b})$ is bounded uniformly in N. Indeed,

$$E||b - \bar{b}||^2 = E(||b|| - \sqrt{N\bar{q}})^2 \leq \frac{1}{\bar{q}N}E(||b||^2 - N\bar{q})^2 \tag{27}$$

$$= \frac{1}{\bar{q}}E\Omega'(\xi_{12}^{N-1}\xi_{34}^{N-1}) + O(1/N) = O(1), \tag{28}$$

thanks to Eq. (9). As an example, let us examine in detail the first term in (26), that is,

$$i\sqrt{N}E\,E_g(\tanh^2 X)\,\Omega'\exp\left(i\,u'\xi^{N-1}\right) - i\,\bar{q}\sqrt{N}E\,\Omega'\exp\left(i\,u'\xi^{N-1}\right). \tag{29}$$

By a simple second order Taylor expansion and an integration by parts on the Gaussian noise g, one finds

$$E_g \tanh^2 X = E_g \tanh^2 \bar{X} + \frac{t}{N}(b - \bar{b})\bar{b}\,E_g\,\partial_x^2 \tanh^2 x\big|_{x=\bar{X}} \tag{30}$$

$$+\frac{t}{2N}E_g\,\partial_x^2 \tanh^2 x\big|_{x=\bar{X}+\theta(X-\bar{X})}\left(g(b - \bar{b})\right)^2 \tag{31}$$

$$= \bar{q} + \frac{t}{2N}(b - \bar{b})(b + \bar{b})E_g\,\partial_x^2 \tanh^2 \bar{X} \tag{32}$$

$$+\frac{t}{2N}||b - \bar{b}||^2 E_g\left(\partial_x^2 \tanh^2 x\big|_{x=\bar{X}+\theta(X-\bar{X})} - \partial_x^2 \tanh^2 \bar{X}\right) \tag{33}$$

$$+\frac{t^2}{2N^2}E_g\,\partial_x^4 \tanh^2 x\big|_{x=\bar{X}+\theta(X-\bar{X})}[(b - \bar{b})(\bar{b} + \theta(b - \bar{b}))]^2 \tag{34}$$

where $0 \leq \theta \leq 1$. Let analyse each term separately. Recalling the definitions of b and $\bar{b}$, the second term in (32) equals

$$\frac{t}{2N}\Omega'(\eta^{s+1}\eta^{s+2} - N\bar{q})\int d\mu(z)\partial_x^2 \tanh^2(\beta h + z\sqrt{t\,\bar{q} + x}) \tag{35}$$

$$= \frac{t}{\sqrt{N}}\Omega'(\xi_{s+1,s+2}^{N-1})(3Y_0 + 2\bar{q} - 2) + O(1/N), \tag{36}$$

where Y_0 was defined in (13). Another application of Taylor's expansion and integration by parts, together with Cauchy-Schwarz inequality and the fact that the derivatives of the function $\tanh^2(x)$ are bounded, shows that the terms (33) and (34) can be bounded by

$$\frac{k}{N}||b - \bar{b}||^2.$$

Therefore, using the estimate (28), the expression (29) reduces to

$$i\,t(3Y_0 + 2\bar{q} - 2)E\,\Omega'\left[\xi_{s+1,s+2}\exp\left(i\,u'\,\xi^{N-1}\right)\right] + O(N^{-1/2}),$$

and

$$i\sqrt{N}E\left\{(\tanh^2 X - \bar{q})\Omega'\exp\left(i\,u'\xi^{N-1}\right)\right\} = t(2\bar{q} - 2 + 3Y_0)\psi' + o(1),$$

where

$$\psi' \equiv \psi_{N-1}^{t'}(u').$$

The other terms in (26) are much simpler than (29), and can be dealt with in the same way. Finally, the whole expression (26) can be rewritten as

$$
t(2\bar{q} - 2 + 3Y_0)\,\psi' - u_{rr'}(1 - \bar{q}^2)\,\phi' \tag{37}
$$
$$
-(\bar{q} - \bar{q}^2)\sum_{a \neq r,r'}(u_{ar} + u_{ar'})\phi' - (Y_0 - (1 - \bar{q})^2)\sum_{\substack{(c,d)\,c,d \neq r,r'}}u_{cd}\,\phi' + o(1).
$$

The steps leading to expression (37) can be repeated with minor changes for the remaining terms (20) to (22). These terms, although they look more complicated than (19) at first sight, are actually simpler to treat, since a first (instead of second) order Taylor expansion around $X = \bar{X}$ is sufficient. This is due to the presence of terms like $\dot{\eta}_a\dot{\eta}_b/N$ or $\dot{\eta}_a b/N$, which are with large probability small, thanks to (9). Also in this case, one finds that terms (20) to (22) give quantities linear in $\phi', \partial\phi', \varphi', \psi'$, apart from terms of order $o(1)$. As for the "error term" S which appears in Theorem 4, one can easily check that it vanishes in the thermodynamic limit. This is a consequence of the exponential decay of overlap fluctuations, as expressed by (15).

Next, we show that terms like $\phi_{N-1}^{t'}(u')$ or $\psi_{N-1}^{t'}(u')$ can be substituted by the same functions calculated at N, t, u, apart from negligible error terms. Indeed, for instance,

$$
\begin{aligned}
\phi_N^t(u) &= \left\langle\exp\left(i\,u'\,\xi^{N-1} + i\,u\,(\sigma_N^1\sigma_N^2 - \bar{q})/\sqrt{N}\right)\right\rangle_t \\
&= \left\langle\exp i\,u'\,\xi^{N-1}\right\rangle_t(1 + o(1)) = \phi' + o(1).
\end{aligned}
$$

In the last step, we used Theorem 4 in order to substitute t with t'. Therefore, Eq. (23) reduces to a linear relation between ϕ, φ and ψ, apart from a remainder which becomes irrelevant in the thermodynamic limit. In the same way, one sees that also Eqs. (24), (25) yield linear equations for ϕ, φ, ψ. Putting everything together, in the thermodynamic limit one has a set of coupled linear differential equations of the form

$$\mathbf{\Phi}^t(u) = \phi^t(u)\,\mathbf{v}(u) + t\,\hat{M}\,\mathbf{\Phi}^t(u) \tag{38}$$

where $\boldsymbol{\Phi}^t(u)$ is the vector

$$\boldsymbol{\Phi}^t(u) = (\partial_{u_{12}}\phi^t(u), \ldots, \partial_{u_{s-1,s}}\phi^t(u), \varphi^{(1)\,t}(u), \ldots, \varphi^{(s)\,t}(u), \psi^t(u)).$$

$\mathbf{v}(u)$ is a vector whose components are homogeneous linear functions of the variables u, while $\hat{M}$ is a real square matrix with elements depending on $\bar{q}, Y_0$ alone. We do not report here the explicit expressions of $\mathbf{v}(u)$ and $\hat{M}$, which are quite complicated. However, it is instructive to check that, for instance, the term (37) is in agreement with this structure. In fact, the coefficient of ϕ' is a homogeneous linear function of the u variables, while the coefficient of ψ' is linear in t and depends only on Y_0 and $\bar{q}$. As will be clear in the following, only the structure (38), and not the specific form of $\mathbf{v}$ and $\hat{M}$, are needed to conclude the proof of the theorem.

Assume at first that the matrix $(1 - t\,\hat{M})$ is invertible, which in principle can fail only for a finite number of values of t, since $\hat{M}$ is finite dimensional. In this case, Eq. (38) can be reduced to a first order differential system in normal form:

$$\boldsymbol{\Phi}^t(u) = \phi^t(u)(1 - t\,\hat{M})^{-1}\mathbf{v}(u), \tag{39}$$

which can be easily integrated. The most general solution for $\phi^t(u)$, compatible with the initial condition

$$\phi^t(0) = 1,$$

is of the form

$$\phi^t(u) = \exp\left\{-\frac{1}{2}(\hat{K}u, u) + (p, u)\right\}, \tag{40}$$

where p is some $s(s-1)/2$ dimensional u-independent vector, and $\hat{K}$ is a $s(s-1)/2 \times s(s-1)/2$ real symmetric positive definite matrix. The symmetry and non negativity of $\hat{K}$ derive from the obvious property of symmetry among replicas, and from the bound

$$\left|\phi^t(u)\right| \leq 1,$$

which holds for any characteristic function. The quadratic dependence on u of the exponent of $\phi^t(u)$ stems from the linear dependence of the components of $\mathbf{v}(u)$. Clearly, Eq. (40) means that the random variables $\{\xi_{ab}^N\}$ converge to some Gaussian process $\{\xi_{ab}\}$. Moreover, it turns out that the identification

$$p = 0$$

and

$$\hat{K} = \hat{L}$$

are straightforward. Indeed, it was shown by Guerra in [17] that, if the limit process is Gaussian, then it is centered and its covariance function is exactly $\hat{L}$.

In order to conclude the proof, it remains to show convergence of the characteristic function for those possible values $\tilde{t}$ where $(1 - t\,\hat{M})$ is singular. For any $\delta > 0$ one can write

$$\phi_N^{\tilde{t}}(u) = \phi_N^{\tilde{t}-\delta}(u) + \delta\,\partial_t\,\phi_N^t\big|_{t=\tilde{t}-\theta_N\,\delta}\,,$$

where $0 < \theta_N < 1$. After a straightforward computation one finds that

$$\partial_t\phi_N^t = \frac{1}{2}\left\langle e^{i\,u\xi^N}\left(\sum_{(a,b)}(\xi_{ab}^N)^2 - s\sum_{a=1}^{s}(\xi_{a,s+1}^N)^2 + \frac{s(s+1)}{2}(\xi_{s+1,s+2}^N)^2\right)\right\rangle.$$

By exploiting the uniform bound (9) and the arbitrariness of δ, one finds therefore that the theorem holds also for $t = \tilde{t}$. $\square$

7   Fluctuations of the free energy

In order to prove Theorem 2, we show that the characteristic function of $\hat{f}_N$ converges to that of $\mathcal{N}(0, \sigma^2(t, x, h))$, i.e.,

$$\lim_{N\to\infty} E\,e^{i\,u\,\hat{f}_N(t,x,h)} = e^{-\frac{u^2}{2}\sigma^2(t,x,h)}.$$

Define

$$\bar{\alpha}(t') = \bar{\alpha}(t', x_{t'}, h)$$
$$\zeta_N(t') = \frac{\ln Z_N(t', x_{t'}, h; J)}{N},$$

where $x_{t'}$ is defined in Eq. (5). The characteristic function of $\hat{f}_N$ can be written as

$$E\,e^{i\,u\hat{f}_N(t,x,h)} \;=\; E\,e^{i\,u\hat{f}_N(0,x_0,h)} + i\,u\,E\int_0^t e^{i\,u\hat{f}_N(t',x_{t'},h)}\frac{d}{dt'}\hat{f}_N(t', x_{t'}, h)dt'. \tag{41}$$

Since

$$\frac{d}{dt'}\bar{\alpha}(t') = \frac{1}{4}(1-\bar{q})^2$$
$$\frac{d}{dt'}\zeta_N(t') = \frac{1}{2\sqrt{t'}N^{3/2}}\sum_{(i,j)}J_{ij}\Omega_{t'}(\sigma_i\sigma_j) - \frac{\bar{q}}{2N\sqrt{x_{t'}}}\sum_i J_i\Omega_{t'}(\sigma_i),$$

one finds through integration by parts that

$$E\left\{e^{i\,u\,\zeta_N(t')}\frac{d}{dt'}\zeta_N(t')\right\} \;=\; \frac{1}{4}E\left\{e^{i\,u\,\zeta_N(t')}[1 - \Omega_{t'}(q_{12}^2) - 2\bar{q}\,(1-\Omega_{t'}(q_{12}))]\right\} \tag{42}$$

$$+\frac{i\,u}{4N}E\left\{e^{i\,u\,\zeta_N(t')}[\Omega_{t'}(q_{12}^2) - 2\bar{q}\,\Omega_{t'}(q_{12}) - N^{-1}]\right\}.$$

14

By using (42) in Eq. (41), one finds

$$
\begin{aligned}
E\,e^{i\,u\hat{f}_N(t,x,h)} \;=\;& E\,e^{i\,u\hat{f}_N(0,x_0,h)} + \frac{u^2\bar{q}^2}{4}E\int_0^t e^{i\,u\hat{f}_N(t')}\,dt' \qquad (43)\\
& -\frac{u^2}{4N}\int_0^t E\,e^{i\,u\hat{f}_N(t')}\left(\Omega_{t'}(\xi_{12}^2)-1\right)\,dt'\\
& -\frac{i\,u}{4\sqrt{N}}\int_0^t E\,e^{i\,u\hat{f}_N(t')}\Omega_{t'}(\xi_{12}^2)\,dt'.
\end{aligned}
$$

At $t=0$, all sites are decoupled and the central limit theorem for i.i.d. random variables implies that

$$
\hat{f}_N(0,x,h) \xrightarrow{d} \mathcal{N}(0,\sigma^2(0,x,h)). \qquad (44)
$$

The last two terms in Eq. (43) clearly vanish for $N\to\infty$. For instance,

$$
N^{-\frac{1}{2}}\left|E\,e^{i\,u\hat{f}_N(t')}\Omega_{t'}(\xi_{12}^2)\right| \le N^{-\frac{1}{2}}E\Omega_{t'}(\xi_{12}^2) = O(N^{-\frac{1}{2}}),
$$

since

$$
E\Omega_{t'}(\xi_{12}^2) = O(1)
$$

for $t < t_c$. Therefore, Eq. (43) yields the following linear integral equation for the characteristic function:

$$
E\,e^{i\,u\hat{f}_N(t,x,h)} = E\,e^{i\,u\hat{f}_N(0,x_0,h)} + \frac{u^2\bar{q}^2}{4}E\int_0^t e^{i\,u\hat{f}_N(t',x_{t'},h)}\,dt' + o(1),
$$

whose solution is, keeping into account the initial condition (44),

$$
E\,e^{i\,u\hat{f}_N(t,x,h)} = e^{-\frac{u^2}{2}\sigma^2(t,x,h)} + o(1).
$$

□

Before concluding this Section, we wish to note that from Eq. (43) one can also obtain in a very simple way a well known result for free energy fluctuations at zero external field and $t < 1$ [4, 6], i.e.,

$$
\eta_N^t \equiv \ln Z_N(t) - N\left(\ln 2 + \frac{t}{4}\right) \xrightarrow{d} \hat{Y}_t - \frac{1}{4}\ln\frac{1}{1-t}, \qquad (45)
$$

where Y_t is a centered Gaussian random variable of variance

$$
\frac{1}{2}\left(\ln\frac{1}{1-t} - t\right).
$$

Indeed, setting $u = \sqrt{N}s$ and $x = h = 0$ in Eq. (43), one obtains the equation

$$
E\,e^{i\,s\,\eta_N^t} = 1 - \frac{u^2}{4}\int_0^t E\,e^{i\,s\,\eta_N^{t'}}\left(\Omega_{t'}(\xi_{12}^2)-1\right)\,dt' - \frac{iu}{4}\int_0^t E\,e^{i\,s\,\eta_N^{t'}}\Omega_{t'}(\xi_{12}^2)\,dt'. \quad (46)
$$

Since Theorem 1 implies, for vanishing external field and $t < 1$,

$$E \left(\Omega_t(\xi_{12}^2) - \langle \xi_{12}^2 \rangle \right)^2 = \langle \xi_{12}^2 \, \xi_{34}^2 \rangle - \langle \xi_{12}^2 \rangle^2 = o(1),$$

Eq. (46) yields

$$E \, e^{i \, s \, \eta_N^t} = 1 - \frac{u^2}{4} \int_0^t E \, e^{i \, s \, \eta_N^{t'}} \left(\frac{1}{1 - t'} - 1 \right) \, dt' - \frac{iu}{4} \int_0^t E \, e^{i \, s \, \eta_N^{t'}} \frac{1}{1 - t'} \, dt' + o(1),$$

from which the result (45) easily follows.

8   Conclusions and outlook

We have employed the cavity method to prove a central limit theorem for the fluctuations of overlaps and free energy, in a region above the Almeida-Thouless line. The key ingredient was provided by the control of the coupled two replica system. The open question remains to understand whether and how our method can be extended to the entire physically expected high temperature region.

In the case of vanishing external field, our method can be employed to obtain very detailed informations on the system in proximity of the critical point. In particular, one can obtain lower and upper bounds on the overlap fluctuations, at $\beta = 1$. We plan to report soon on this [24].

References

[1] D. Sherrington, S. Kirkpatrick, *Solvable model of a spin-glass*, Phys. Rev. Lett. **35**, 1792 (1975).

[2] S. Kirkpatrick, D. Sherrington, *Infinite-ranged models of spin-glasses*, Phys. Rev. B **17**, 4384 (1978).

[3] M. Mézard, G. Parisi and M. A. Virasoro, *Spin glass theory and beyond*, World Scientific, Singapore (1987).

[4] M. Aizenman, J. Lebowitz and D. Ruelle, *Some rigorous results on the Sherrington-Kirkpatrick spin glass model*, Commun. Math. Phys. **112**, 3 (1987).

[5] F. Guerra, *Fluctuations and thermodynamic variables in mean field spin glass models*, in: *Stochastic processes, physics and geometry, II*, S. Albeverio *et al.*, eds., Singapore (1995).

[6] F. Comets, J. Neveu, *The Sherrington-Kirkpatrick model of spin glasses and stochastic calculus: the high temperature case*, Commun. Math. Phys. **166**, 549 (1995).

[7] J. Fröhlich, B. Zegarlinski, *Some comments on the Sherrington-Kirkpatrick model of spin glasses*, Commun. Math. Phys. **112**, 553 (1987).

[8] M. Talagrand, *On the high temperature region of the Sherrington-Kirkpatrick model*, C. R. A. S., to appear; M. Talagrand, *On the high temperature phase of the Sherrington-Kirkpatrick model*, Ann. Probab., to appear.

[9] M. Talagrand, *Mean field models for spin glasses: a first course*, course given in Saint Flour in summer 2000, to appear.

[10] M. Talagrand, *Replica symmetry breaking and exponential inequalities for the Sherrington Kirkpatrick model*, Ann. Probab. **28**, 1018 (2000).

[11] L. Pastur, M. Shcherbina, *The absence of self-averaging of the order parameter in the Sherrington-Kirkpatrick model*, J. Stat. Phys. **62**, 1 (1991).

[12] F. Guerra, *About the overlap distribution in mean field spin glass models*, Int. Jou. Mod. Phys. B **10**, 1675 (1996).

[13] S. Ghirlanda, F. Guerra, *General properties of overlap distributions in disordered spin systems. Towards Parisi ultrametricity*, J. Phys. A **31**, 9149 (1998).

[14] F. Guerra, F. L. Toninelli, *Quadratic replica coupling for the Sherrington-Kirkpatrick mean field spin glass model*, J. Math. Phys., **43**, 3704 (2002).

[15] A. Bovier, I. Kurkova, M. Löwe, *Fluctuations of the free energy in the REM and the p-spin SK models*, Ann. Probab., to appear.

[16] M. Talagrand, *Spin glasses: a challenge for mathematicians. Mean field models and cavity method*, Springer-Verlag, to appear.

[17] F. Guerra, *Sum rules for the free energy in the mean field spin glass model*, Fields Institute Communications **30**, 161 (2001).

[18] F. Guerra, *The cavity method in the mean field spin glass model. Functional representations of thermodynamic variables*, in: *Advances in dynamical systems and quantum physics*, S. Albeverio *et al.*, eds., Singapore (1995).

[19] M. Aizenman, P. Contucci, *On the stability of the quenched state in mean field spin glass models*, J. Stat. Phys. **92**, 765 (1998).

[20] G. Parisi, *On the probabilistic formulation of the replica approach to spin glasses*, cond-mat/9801081.

[21] M. Talagrand *A new look at independence*, Ann. Probab. **24**, 1 (1996).

[22] M. Talagrand, *Concentration of measure and isoperimetric inequalities in product spaces*, Publ. Math. I.H.E.S. **81**, 73 (1995).

[23] A. N. Shiryaev, *Probability*, Springer, Berlin (1989).

[24] F. Guerra, F. L. Toninelli, in preparation